E. Slobodyansky

#5

RAW SEQUENCE LISTING
PATENT APPLICATION *US/09/415,540*

DATE: 11/09/1999
TIME: 15:46:25

INPUT SET: S33887.raw

This Raw Listing contains the General Information Section and up to the first 5 pages.

ENTERED

SEQUENCE LISTING

(1) General Information

(i) APPLICANT: Hawkins, Phillip R.
Hillman, Jennifer L.

(ii) TITLE OF THE INVENTION: A NOVEL HUMAN PYROPHOSPHATASE

(iii) NUMBER OF SEQUENCES: 5

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Incyte Pharmaceuticals, Inc.
(B) STREET: 3174 Porter Drive
(C) CITY: Palo Alto
(D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94304

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: DOS
(D) SOFTWARE: FastSEQ Version 1.5

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/415,540
(B) FILING DATE:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/741,437
(B) FILING DATE:

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Billings, Lucy J.
(B) REGISTRATION NUMBER: 36,749
(C) REFERENCE/DOCKET NUMBER: PF-0148 US

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (415) 855-0555
(B) TELEFAX: (415) 845-4166

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 289 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Ser Gly Phe Ser Thr Glu Glu Arg Ala Ala Pro Phe Ser Leu Glu
 1               5                  10                  15
Tyr Arg Val Phe Leu Lys Asn Glu Lys Gly Gln Tyr Ile Ser Pro Phe
            20                  25                  30
His Asp Ile Pro Ile Tyr Ala Asp Lys Asp Val Phe His Met Val Val
        35                  40                  45
Glu Val Pro Arg Trp Ser Asn Ala Lys Met Glu Ile Ala Thr Lys Asp
    50                  55                  60
Pro Leu Asn Pro Ile Lys Gln Asp Val Lys Lys Gly Lys Leu Arg Tyr
65                  70                  75                  80
Val Ala Asn Leu Phe Pro Tyr Lys Gly Tyr Ile Trp Asn Tyr Gly Ala
                85                  90                  95
Ile Pro Gln Thr Trp Glu Asp Pro Gly His Asn Asp Lys His Thr Gly
            100                 105                 110
Cys Cys Gly Asp Asn Asp Pro Ile Asp Val Cys Glu Ile Gly Ser Lys
        115                 120                 125
Val Cys Ala Arg Gly Glu Ile Ile Gly Val Lys Val Leu Gly Ile Leu
    130                 135                 140
Ala Met Ile Asp Glu Gly Glu Thr Asp Trp Lys Val Ile Ala Ile Asn
145                 150                 155                 160
Val Asp Asp Pro Asp Ala Ala Asn Tyr Asn Asp Ile Asn Asp Val Lys
                165                 170                 175
Arg Leu Lys Pro Gly Tyr Leu Glu Ala Thr Val Asp Trp Phe Arg Arg
            180                 185                 190
Tyr Lys Val Pro Asp Gly Lys Pro Glu Asn Glu Phe Ala Phe Asn Ala
        195                 200                 205
Glu Phe Lys Asp Lys Asp Phe Ala Ile Asp Ile Ile Lys Ser Thr His
    210                 215                 220
Asp His Trp Lys Ala Leu Val Thr Lys Lys Thr Asn Gly Lys Gly Ile
225                 230                 235                 240
Ser Cys Met Asn Thr Thr Leu Ser Glu Ser Pro Phe Lys Cys Asp Pro
                245                 250                 255
Asp Ala Ala Arg Ala Ile Val Asp Ala Leu Pro Pro Pro Cys Glu Ser
            260                 265                 270
Ala Cys Thr Val Pro Thr Asp Val Asp Lys Trp Phe His His Gln Lys
        275                 280                 285
Asn
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1275 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vii) IMMEDIATE SOURCE:
(A) LIBRARY:
(B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

CAAGAGGTTN GGGGCTCTCT CCTTGTCAGT CGGCGCCGCG TGCGGGCTGG TGGCTCTGTG 60
GCAGCGGCGG CGGCAGGACT CCGGCACTAT GAGCGGCTTC AGCACCGAGG AGCGCGCCGC 120
GCCCTTCTCC CTGGAGTACC GAGTCTTCCT CAAAAATGAG AAAGGACAAT ATATATCTCC 180
ATTTCATGAT ATTCCAATTT ATGCAGATAA GGATGTGTTT CACATGGTAG TTGAAGTACC 240
ACGCTGGTCT AATGCAAAAA TGGAGATTGC TACAAAGGAC CCTTTAAACC CTATTAAACA 300
AGATGTGAAA AAAGGAAAAC TTCGCTATGT TGCGAATTTG TTCCCGTATA AAGGATATAT 360
CTGGAACTAT GGTGCCATCC CTCAGACTTG GGAAGACCCA GGGCACAATG ATAAACATAC 420
TGGCTGTTGT GGTGACAATG ACCCAATTGA TGTGTGTGAA ATTGGAAGCA AGGTATGTGC 480
AAGAGGTGAA ATAATTGGCG TGAAAGTTCT AGGCATATTG GCTATGATTG ACGAAGGGGA 540
AACCGACTGG AAAGTCATTG CCATTAATGT GGATGATCCT GATGCAGCCA ATTATAATGA 600
TATCAATGAT GTCAAACGGC TGAAACCTGG CTACTTAGAA GCTACTGTGG ACTGGTTTAG 660
AAGGTATAAG GTTCCTGATG GAAAACCAGA AAATGAGTTT GCGTTTAATG CAGAATTTAA 720
AGATAAGGAC TTTGCCATTG ATATTATTAA AAGCACTCAT GACCATTGGA AAGCATTAGT 780
GACTAAGAAA ACGAATGGAA AAGGAATCAG TTGCATGAAT ACAACTTTGT CTGAGAGCCC 840
CTTCAAGTGT GATCCTGATG CTGCCAGAGC CATTGTGGAT GCTTTACCAC CACCCTGTGA 900
ATCTGCCTGC ACAGTACCAA CAGACGTGGA TAAGTGGTTC CATCACCAGA AAAACTAATG 960
AGATTTCTCT GGAATACAAG CTGATATTGC TACATCGTGT TCATCTGGAT GTATTAGAAG 1020
TAAAAGTAGT AGCTTTTCAA AGCTTTAAAT TTGTAGAACT CATCTAACTA AAGTAAATTC 1080
TGCTGTGACT AATCCAATAT ACTCAGAATG TTATCCATCT AAAGCATTTT TCATATCTCA 1140
ACTAAGATAA CTTTTAGCAC ATGCTTAAAT ATCAAAGCAG TTGTCATTTG GAAGTCACTT 1200
GTGAATAGAT GTGCAAGGGG AGCACATATT GGATGTATAT GTTACCATAT GTTAGGAAAT 1260
AAAATTATTT TGCTG 1275

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 114 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 727225

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
153
154   Asn Ile Phe Pro Tyr Lys Gly Tyr Ile Trp Asn Tyr Gly Thr Leu Pro
155    1               5                  10                  15
156   Gln Thr Trp Glu Asp Pro His Glu Lys Asp Lys Ser Thr Asn Cys Phe
157               20                  25                  30
158   Gly Asp Asn Asp Pro Ile Asp Val Cys Glu Ile Gly Ser Lys Ile Leu
159           35                  40                  45
160   Ser Cys Gly Glu Val Ile His Val Lys Ile Leu Gly Ile Leu Ala Leu
161       50                  55                  60
162   Ile Asp Glu Gly Glu Thr Asp Trp Lys Leu Ile Ala Ile Asn Ala Asn
163   65                  70                  75                  80
164   Asp Pro Glu Ala Ser Lys Phe His Asp Ile Asp Asp Val Lys Lys Phe
165                   85                  90                  95
166   Lys Pro Gly Tyr Leu Glu Ala Thr Leu Asn Trp Phe Arg Leu Tyr Lys
167               100                 105                 110
168   Val Pro
169
170
171            (2) INFORMATION FOR SEQ ID NO:4:
172
173          (i) SEQUENCE CHARACTERISTICS:
174            (A) LENGTH: 289 amino acids
175            (B) TYPE: amino acid
176            (C) STRANDEDNESS: single
177            (D) TOPOLOGY: linear
178
179          (ii) MOLECULE TYPE: peptide
180
181          (vii) IMMEDIATE SOURCE:
182             (A) LIBRARY: GenBank
183             (B) CLONE: 585322
184
185          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
186
187   Met Ser Ser Phe Ser Ser Glu Glu Arg Ala Ala Pro Phe Thr Leu Glu
188    1               5                  10                  15
189   Tyr Arg Val Phe Leu Lys Asn Glu Lys Gly Gln Tyr Ile Ser Pro Phe
190               20                  25                  30
191   His Asp Ile Pro Ile Tyr Ala Asp Lys Glu Val Phe His Met Val Val
192           35                  40                  45
193   Glu Val Pro Arg Trp Ser Asn Ala Lys Met Glu Ile Ala Thr Lys Asp
194       50                  55                  60
195   Pro Leu Asn Pro Ile Lys Gln Asp Val Lys Lys Gly Lys Leu Arg Tyr
196   65                  70                  75                  80
197   Val Ala Asn Leu Phe Pro Tyr Lys Gly Tyr Ile Trp Asn Tyr Gly Ala
198                   85                  90                  95
199   Ile Pro Gln Thr Trp Glu Asp Pro Gly His Asn Asp Lys His Thr Gly
200               100                 105                 110
201   Cys Cys Gly Asp Asn Asp Pro Ile Asp Val Cys Glu Ile Gly Ser Lys
202           115                 120                 125
203   Val Cys Ala Arg Gly Glu Ile Ile Arg Val Lys Val Leu Gly Ile Leu
204       130                 135                 140
205   Ala Met Ile Asp Glu Gly Glu Thr Asp Trp Lys Val Ile Ala Ile Asn
```

```
145                 150                 155                 160
Val Glu Asp Pro Asp Ala Ala Asn Tyr Asn Asp Ile Asn Asp Val Lys
                165                 170                 175
Arg Leu Lys Pro Gly Tyr Leu Glu Ala Thr Val Asp Trp Phe Arg Arg
            180                 185                 190
Tyr Lys Val Pro Asp Gly Lys Pro Glu Asn Glu Phe Ala Phe Asn Ala
        195                 200                 205
Glu Phe Lys Asp Lys Asn Phe Ala Ile Asp Ile Ile Glu Ser Thr His
    210                 215                 220
Asp Tyr Trp Arg Ala Leu Val Thr Lys Lys Thr Asp Gly Lys Gly Ile
225                 230                 235                 240
Ser Cys Met Asn Thr Thr Val Ser Glu Ser Pro Phe Gln Cys Asp Pro
                245                 250                 255
Asp Ala Ala Lys Ala Ile Val Asp Ala Leu Pro Pro Pro Cys Glu Ser
            260                 265                 270
Ala Cys Thr Ile Pro Thr Asp Val Asp Lys Trp Phe His His Gln Lys
        275                 280                 285
Asn
```

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 287 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide
(vii) IMMEDIATE SOURCE:
(A) LIBRARY: GenBank
(B) CLONE: 4199

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
Met Thr Tyr Thr Thr Arg Gln Ile Gly Ala Lys Asn Thr Leu Glu Tyr
 1               5                  10                  15
Lys Val Tyr Ile Glu Lys Asp Gly Lys Pro Val Ser Ala Phe His Asp
            20                  25                  30
Ile Pro Leu Tyr Ala Asp Lys Glu Asn Asn Ile Phe Asn Met Val Val
        35                  40                  45
Glu Ile Pro Arg Trp Thr Asn Ala Lys Leu Glu Ile Thr Lys Glu Glu
    50                  55                  60
Thr Leu Asn Pro Ile Ile Gln Asp Thr Lys Lys Gly Lys Leu Arg Phe
65                  70                  75                  80
Val Arg Asn Cys Phe Pro His His Gly Tyr Ile His Asn Tyr Gly Ala
                85                  90                  95
Phe Pro Gln Thr Trp Glu Asp Pro Asn Val Ser His Pro Glu Thr Lys
            100                 105                 110
Ala Val Gly Asp Asn Asp Pro Ile Asp Val Leu Glu Ile Gly Glu Thr
        115                 120                 125
Ile Ala Tyr Thr Gly Gln Val Lys Gln Val Lys Ala Leu Gly Ile Met
    130                 135                 140
```

SEQUENCE CORRECTION REPORT
PATENT APPLICATION *US/09/415,540*

DATE: 11/09/1999
TIME: 15:46:27

INPUT SET: S33887.raw

Line	Original Text	Corrected Text
3	(1) General Information	(1) GENERAL INFORMATION:
8	(ii) TITLE OF THE INVENTION: A NOVEL HUMAN PY	(ii) TITLE OF INVENTION: A NOVEL HUMAN PYROP